

Location, location, location: Challenges of Outsourced Usability Evaluation

John Murphy¹, Steve Howard², Jesper Kjeldskov³ and Steve Goschnick²

¹ Design4Use
7 Abbotsford Street
Melbourne, Victoria 3067
Australia
john@design4use.com.au

² Department of Information Systems
The University of Melbourne
Parkville, Victoria 3010
Australia
{showard, stevenbg}@unimelb.edu.au

³ Department of Computer Science
Aalborg University
DK-9220 Aalborg East
Denmark
jesper@cs.auc.dk

ABSTRACT

This position paper presents some of the challenges experienced in relation to an outsourced usability evaluation of commercial collaboration product, which we would like to raise in the Improving the Interplay between Usability Evaluation and User Interface Design workshop. The paper describes the context of the outsourced evaluation, three challenges of location, how the evaluation was carried out and reported. Finally, we outline some of the lessons learned.

INTRODUCTION

A commercial company is developing a new product, which is intended to support collaborative work amongst non-technical commercial workers. For this product to succeed, non-technical users must be able to use the tool easily. A significant component of the ease of use of the product is the users' ability to create a clear and coherent mental model of the system. In order to evaluate the current design of the product, it was decided to conduct a usability evaluation of the current design. The overall objective for the usability test was to determine whether the product supports a coherent and consistent mental model for a user collaboratively sharing files with others to achieve a goal. The secondary goal of the evaluation was to determine whether the interface screen design and flow supports the individual tasks of creating and sharing through the product with another person and accepting an invitation to share.

The evaluation allowed us to study some of the challenges of outsourcing usability in a large industrial software development project. In the following sections, we first briefly introduce the context of the product usability evaluation. Secondly, we outline some of the challenges encountered in planning and conducting the evaluation, which we would like to address in the workshop. Hereafter, we briefly describe how the evaluation was carried out and the mechanisms employed for reporting them. Finally, we outline some of the lessons learned.

BACKGROUND

The product is being developed within a multi-national software product company based in the United States. Typical of this type of company, the product company has a multitude of existing and new products under development in various programs under aggressive time and resource constraints. The company has a strong commitment to being focused on the needs of customers in relation to their products and services. As such, the company has strong human computer interaction (HCI) skills supporting the development of user interfaces that are easy to use. However, the number of these resources is limited in relation to the number of projects and amount of HCI work required. As with many companies throughout the world, this product company is investigating an outsourcing model to support HCI requirements and in particular usability evaluation.

The company has offices in Australia that, aside from day-to-day business are involved in HCI based research in collaboration with the Universities of Melbourne and Aalborg. This program has been running for over four years encompassing collaboration on developing research techniques, industry projects, teaching and sponsorship of a state of the art usability laboratory in The University of Melbourne, Department of Information Systems. The university has strong skills and resources in usability evaluation and is very active in research and teaching of evaluation techniques.

Given this collaborative relationship, the company decided to pilot a set of usability tests on one of their developing products using the Company – Melbourne – Aalborg relationship. Following discussions with senior company product managers, the product was selected as a suitable candidate based on it being at an appropriate state of development, requiring HCI support and being an open source development which circumvented non-disclosure requirements.

From the company perspective, the objective of the testing was to determine whether cost effective useful findings

could be established through timely testing (as discussed in e.g. Kjeldskov et al. 2004). This required timely setting up of the software, designing the test, recruiting participants, running the sessions and analyzing and reporting of results. The design of the testing had to be determined appropriately with the knowledge that development was continuing throughout the testing period and the testing should be budgeted to be cost effective relative to running the testing in the United States.

CHALLENGES TO THE EVALUATION OF THE PRODUCT

Usability testing and evaluation faces challenges: some generic and some features of the particularities of the evaluation under question; some interesting and others mundane. In this section we focus on three challenges that we found particularly problematic: location, location and location.

Location – Geography

Conducting a remote usability evaluation places a particular burden on communication and the maintenance of situation awareness (Murphy 2001, Hartson et al. 1996). Multiplexed time zones can aid in rapid turn around of results but only if synchronous interaction is not required at times of unavailability, or indeed uncivilised hours, and only if the disparate teams are ‘talking the same language’.

Prior to commencing the evaluation, and drawing on a mix of local knowledge, documentation, email and teleconferencing skills, we harvested as much understanding of the remote situation as we were able. In the workshop we will discuss the influence that the following had over the project:

- Expectations on rapid turn around time and streamlined reporting requirements
- Preferences for and bias toward different data collection methods and data types
- Concern that usability evaluation produce more than merely a list of problems (i.e. the results should be translated into design change suggestions)
- Interest in the process (how the evaluation was conducted) as opposed to merely the product and the findings from the evaluation.

Location - Sector

Combining multiple sectors (in this case industry practitioners and university researchers and research students) is a real strength of our approach. The established and ongoing relationship between the company and the Universities of Melbourne and Aalborg allows us to respond rapidly to emerging opportunities under the rubric of a tested agreement. However, as a cross sectoral collaboration it is not without its frustrations (but see Lambert, 2003 for some solutions). Some of the issues we will raise in the workshop include:

- The need for industry partners to be able to guarantee short cycle delivery times whilst recognising the imperative that university researchers’ engage is risk oriented longer-term discovery.
- The need for university based researches to balance consulting and applied research with more basic enquiry.
- The importance of exposing PhD students to ‘real world’ projects whilst at the same time limiting unnecessary distractions to their ongoing thesis work.
- The management and protection of intellectual property; both background and created intellectual property of the researchers, the students and the industry partner.
- Gauging the benefits that flow from any collaboration, be they immediate and tangible or more speculative.

Location – Development phase

Usability evaluators, be they located in industry or universities, are unfortunately rather experienced at being introduced too late into the lifecycle to have a major impact on the product. It was therefore rewarding to be invited to comment at a relatively early stage in a product’s development (see Rubin, 1994 for a discussion of the importance of life cycle positioning). However, an opportunity to comment early should not be confused with an occasion for unbridled creativity! Some of the issues we should like to raise in the workshop include:

- Gauging the degrees of freedom available in responding to the identified usability flaws.
- The critical importance of the representational form of any feedback to the design team.
- Balancing a critical perspective on the present design with a constructive account of the next.

Faced with these challenges of outsourced usability evaluation, we designed and conducted an evaluation of the product in collaboration between the company and The University of Melbourne and reported the results back to the development team in the United States. The design of the usability evaluation and the way we reporting back the results are described below.

EVALUATION DESCRIPTION

The product usability evaluation was conducted over two days at a state-of-the-art usability laboratory at The University of Melbourne, Australia. The evaluation was done in a collaborative working environment with real life scenarios and tasks requiring the use of other software such as e-mail client and folder and file manipulation tools. Two independent usability evaluations were conducted; a user-based evaluation and a heuristic walkthrough. These are described in detail below.

User- Based Evaluation

The user-based evaluation was based on the think-aloud protocol, involving three triads of test subject working collaboratively through the product. The test subjects were physically separated from each other and could only collaborate using the product and e-mail.

Each of the three evaluation sessions took approximately one hour and consisted of a collaborative task requiring the three users to share information by creating, sharing and using the product. During the evaluation, the subjects were presented with a scenario and tasks to complete.

The scenario was based on the common financial task of sharing and updating work plans within a finance group. This scenario was selected as common across many companies and performed by staff requiring no particular technical knowledge. The profile of the test subjects were non-technical knowledge workers who, ideally, could be part of a team who are used to working together. The subjects were not employed by the company and did not have any special knowledge of the company software.

The user-based evaluation sessions were recorded on digital video capturing all overviews of all three test subjects and their respective computer monitors.

Heuristic Walkthrough

Secondly, three Doctoral students specializing in Human-Computer Interaction conducted a Heuristic Walkthrough of the product software using the scenarios described above.

The Heuristic Walkthrough session lasted approximately ninety minutes and was facilitated by the first author who recorded usability problems by the expert reviewers for later analysis and comparison with the user based data.

REPORTING THE RESULTS

The evaluation had several audiences - project stakeholders in the form of product managers and senior product development staff, company HCI professionals based in the United States and most importantly, product engineers actually working on the product. Each of the different audiences required different information; the project stakeholders were most concerned with the feasibility of outsourced usability evaluation in terms of costs, resources and overall effectiveness; the HCI professionals were concerned to validate the evaluation process and results to both ensure the quality of the results for the product work, but more importantly to investigate how and whether this process and resource might be able to support on-going company HCI work; and the product engineers wanted "design ready" results. From a product engineering perspective, it was understood that the reporting of problems would not be useful without some accompanying proposal of a solution, particularly in the case of significant or complex problems.

Given these different audiences and reporting requirements, a number of different reporting mechanisms were employed. A telephone conference was used to report high level findings, costings and an overall project feasibility to stakeholders and HCI staff. A short highlights video of the usability laboratory, equipment and 'snippets' of the actual evaluation was prepared to present the evaluation process to the company HCI staff and stakeholders. A written evaluation report was prepared explaining the results in detail for product engineers and company HCI staff. It was structured with a usability problem summary table, a discussion of each of the usability issues, user interface design solution ideas and a description of the test.

The evaluation results were well received by the company in the United States. The cost of running the evaluation was within budget and is believed to be a cost effective opportunity for the company. Further investigating into the outsourcing model is currently in progress.

LESSONS LEARNED

The product software is still under development and prone to errors. These factors led to a significant increase in the standard level of support and intervention required for usability testing. For instance, participants required support where the ability of a user was significantly different to the other team members and needed to maintain timely collaboration with colleagues. In cases where participants acting as team leaders sharing files and becoming entangled in Microsoft file-sharing were assisted back to the product environment to maintain the flow of the task. Also, it was important that users were not distracted and did not spend significant cognitive effort on things such as learning an unfamiliar e-mail client or manipulating folders.

In relation to the process of evaluating the product, a strong background contextual knowledge is essential to ensure the testing is effective. Budgets, timelines for product development intended audience are all used to support the design of the evaluation. Other deeper and more subtle knowledge such as market share for this product, future plans to integrate with other products, main competing products and number and skill of engineers available to work on the product are just a sample of the broader knowledge that is useful in supporting the design of the testing.

The physical setup of hardware and software environment and skilled technical support for a product in development is also a challenge. For example, one of the product requirements was a static IP address which was not able to be obtained in the University environment. The company engineers in Australia spend one full man day and University technical staff spent almost half a day setting up the environment and software. This challenge may also be viewed as an advantage in the enforcement of independence through at all levels based on the remoteness of the testing. Not only are the evaluators and evaluation staff independent, but also the entire technical setup is required

to be independent which may in itself reveal technical system 'bugs'.

The video highlights were found to be extremely valuable as a fast effective mechanism of providing a significant amount of information to the project stakeholders and company HCI staff. The video highlights viewed in conjunction with the teleconference meant that the presentation and ensuing discussion quickly became informed and focused.

The value of this type of work to the researchers is in exposure to real world systems, provision of actual data for their research and provision of money to support their facilities. Real world systems expose researchers to actual problems and systems serving to ground their thinking and research ideas in reality. This also applies to usability laboratory staff who broaden their experience in setting up and running real industry evaluations. Providing commercial confidentiality can be preserved and data can be made suitably anonymous this work can be a source of actual data for research projects. This can be implemented through establishing a suitable commercial reviewing process. Finally and significantly, researchers benefit from the money that is directed towards maintaining the quality of the University evaluation facilities and staff.

ACKNOWLEDGEMENTS

We thank the test subjects who participated in the the product evaluation and the Doctoral students who conducted the heuristic walkthrough; Jeni Paay, Sonja Pedell and Jan Skjetne.

REFERENCES

- Hartson, H., Castillo, J., Kelso, J., Kamler, J., & Neale, W. (1996) Remote Evaluation: The Network as an Extension of the Usability Laboratory. Proceedings of CHI'96. http://www.acm.org/sigchi/chi96/proceedings/papers/Hartson/hrh_txt.htm (accessed 8/26 2004)
- Kjeldskov, J., Skov, M. B. and Stage, J. (2004) Instant Data Analysis: Evaluating Usability in a Day. Proceedings of NordiCHI 2004, Tampere, Finland, ACM, pp. 233-240
- Lambert (2003) Lambert Review of Business-University Collaboration http://www.hm-treasury.gov.uk/media/EA556/lambert_review_final_450.pdf (accessed 8/26 2004)
- Murphy, J. (2001) Modelling 'Designer – Tester – Subject' Relationships in International Usability Testing. IWIPS 2001 Designing for Global Markets 3 July 12-14 Milton Keynes, United Kingdom Editors, Donald L. Day and Lynne Dunckley
- Rubin, J. (1994) Handbook of usability testing: how to plan, design, and conduct effective tests. New York: Wiley.